Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 119 (2015) 464 – 469

Procedia Engineering

www.elsevier.com/locate/procedia

13th Computer Control for Water Industry Conference, CCWI 2015

A cloud-based distributed data collection system for decentralised wastewater treatment plants

P. O'Donovan[ab*], D. Coburn[a], E. Jones[a], L. Hannon[a], M. Glavin[a], D. Mullins[a], E. Clifford[abc]

[a] *College of Engineering and Informatics, NUI Galway*
[b] *IRUSE, NUI Galway*
[c] *Ryan Institute, NUI Galway*

Abstract

Compliance with increasingly stringent legislation is one of the key challenges in the wastewater technology sector in the European Union (EU). Many EU countries have chosen to use decentralised wastewater treatment plants (WWTPs) because of reduced infrastructural requirements. However, achieving compliance of regulatory standards using a decentralised strategy can be difficult and costly due to inefficient operating cost-efficiency, infrequent monitoring capability, as well as the logistical challenges faced by operators travelling between sites. With these problems in mind, this paper presents a distributed information system that was developed as part of an Irish-based research project that was funded by Enterprise Ireland. The system focuses on the acquisition and hosting of data from decentralised WWTP's to facilitate reliable and timely access to data that can be used for monitoring and analysis, without incurring the inherent logistical and technical costs associated with manual data collection methods.

Peer-review under responsibility of the Scientific Committee of CCWI 2015

Keywords: water, ict, information systems, decision support, data analytics, data integration

* Corresponding author.

E-mail address: peter.odonovan@nuigalway.ie

Peer-review under responsibility of the Scientific Committee of CCWI 2015
doi:10.1016/j.proeng.2015.08.865

1 Introduction

The existence of wastewater treatment practices that are not inline with EU legislation and standards can result in inefficient and unreliable plant performance, which in turn can result in increased operating costs, reduced water quality and exposes members to possible EU sanctions. However, meeting the increasing operational and performance standards asserted by EU legislation is becoming more challenging. The Water Framework Directive (WFD) [1] and the STRIVE Directive [2], both emphasise the need for reliable plant performance using effective data management, and capability for continuous monitoring and control.

While these ICT-oriented requirements may appear to be rudimentary, the characterisation of the environment in which they are being enforced must be considered. For example, a decentralised and unmanned WWTP with low-bandwidth, will inevitably find it more difficult to employ a visible and real-time monitoring system to promote awareness of water quality and report on the facilities operating efficiency, when compared to a centralised facility with the appropriate network infrastructure.

Therefore, where decentralised wastewater treatment plants (WWTPs) are used, achieving compliance with regulatory standards can be difficult and/or costly for a number of reasons, including poorer effluent quality, inefficient operating cost-efficiency, infrequent monitoring which can lead to delays if problems occur, as well as the logical cost of operators travelling between various unmanned facilities [3]–[5].

To address these issues, the development of low-cost and robust sensors that are capable of providing real-time feedback that can enable operators to make informed decisions [6], [7], has been highlighted as a crucial technology requirement. At present, online sensor technologies are expensive and are developed independently of WWTP control systems. This results in data being presented in a manner that is too granular and not fit for purpose, which makes it time-consuming for engineers or WWTP managers to process and interpret the data.

In this paper, we present the distributed data collection system architecture that was developed during a research project, which is referred to as ITWat. The overarching objective of ITWat was to develop a suite of innovative sensor technologies that can be used for the monitoring and control of decentralised WWTPs. However, to realise these objectives, a reliable and intelligent method for collecting data from decentralised WWTPs was identified as a key technology enabler. Considering the remote and isolated nature of decentralised WWTPs, a novel and innovative aspect of the distributed system architecture presented in this paper, is its operational resilience to facilities with poor bandwidth and connectivity.

The remainder of this paper is organised as follows - (1) a methodology explaining the process for the followed during the system construction, (2) an explanation of the ITWat system architecture and components, and (3) a conclusion to the paper.

2 Methodology

We employed a three-step methodology for the design and development of the ITWat distributed data collection system. First, we undertook planning and requirements to provide the foundations for future implementation. Second, we chose the technologies and frameworks that would be needed to construct the system. Third, we developed a distributed architecture that could be used as a blueprint for the systems implementation, and then developed a proof-of-concept using the technologies identified in our methodology.

2.1 Phase 1 - Planning and requirements

In the initial phase of our methodology, we focused on establishing the state-of-play in relation to remote monitoring of decentralised WWTPs, highlighting any impediments we encountered, and reporting on the actions we undertook.

Actions

- Reviewed low bandwidth distributed computing solutions for suitability to WWTPs.
- Assessed tools and technologies that are currently used in distributed computing.
- Identified data measurements that can be collected from treatment plants.
- Created a profile-based data collection strategy based on bandwidth availability.

- Developed a software delivery schedule for the ITWat data collection system.

Challenges

- The extreme connectivity characteristics of decentralised WWTPs (i.e. bandwidth constraints and deployment environment), as well as the specific data conventions and structure of the SVI data, meant that we were unable to find an existing solutions that could facilitate data collection.

2.2 Phase 2 – Technologies and frameworks

In the second phase of our methodology we chose the technologies that would be required to implement the data collection system. To facilitate distributed operation, we chose a Service Oriented Architecture (SOA) as the basis for the overall system architecture, whereby the core functionality of the system is exposed via cloud-based services. An SOA approach was chosen as it provides a platform independent, technology agnostic, and modular approach to the systems construction. We felt that these characteristics would serve to make the system easier to extend and maintain over time, as well as supporting interoperability with other systems.

The Microsoft .NET Framework 4.0 was chosen to support the development of bespoke components in the system. The particular aspects of the .NET Framework that we identified as being relevant to the implementation included the object-oriented programming language C#, Windows Communication Foundation (WCF) for messaging (i.e. remote procedure calls) and ASP.NET Model View Controller (MVC) for browser-based applications. Microsoft .NET was chosen in place of similar frameworks, such as Java EE, because of the maturity of its development environment, the high level of integration between its Integrated Development Environment (IDE) and other Microsoft tools and platforms (e.g. SQL Server and Azure Cloud). The main constraint with .NET applications is their inability to execute on operating systems other than Microsoft Windows. However, we consider this an acceptable trade-off given our initial scoping did not highlight this as a requirement, and existing systems operating in WWTPs are predominantly Windows-based.

To facilitate open data exchange between the interacting components in the system, we endeavored to promote open and standard protocols throughout the system. In terms of communication protocols, we chose TCP/HTTP for streaming data between WWTPs and the cloud, and to encode these data streams we chose JSON and SOAP/XML. JSON was chosen to encode numerical measurements because of its unobtrusive and lightweight format, which in turn results in smaller packets when transmitting data to the cloud. In contrast, SOAP/XML was chosen for transferring binary files because JSON encoding of binary data is approximately 40% larger than what can be achieved using SOAP/XML.

2.3 Phase 3 - Implementation of architecture

In the final phase of our methodology, we created a proof-of-concept for the ITWat data collection system, and identified previously hidden constraints or impediments to the implementation. If these obstacles were small, they were incorporated in the next update of the system, while bigger changes were added to our future work list.

Actions

- Initialised and configured the physical components of the system (e.g. cloud server).
- Installed software components of the system (e.g. SQL Server)
- Created software models to describe the data collection process and built high-level architecture.
- Established a naming convention and structure for interfacing with the SVI processing system.
- Constructed bespoke software components to facilitate (a) local data extraction, (b) data transfer over TCP/HTTP, (c) local and cloud-based data persistence, and (d) basic reporting.
- Undertook testing of system using test data and identified initial benchmarks for differing connectivity profiles (e.g. 2G, 3G, ADSL etc.)

Challenges

- There was no convention or standard for interoperating with the SVI processing system.
- Simulating bandwidth constraints that may exist in WWTPs is difficult in a controlled environment that has

high bandwidth – although bandwidth can be limited for the purpose of testing, the ad hoc and intermittent connectivity caused by a poor signal cannot.

3 ITWat system architecture

Figure 1 illustrates the architecture of the ITWat distributed data collection system, with each of its interacting components depicted in the context of their physical location. These components work together to deliver an automated data collection mechanism for numerical measurements and binary files (e.g. image or CSV) from decentralised WWTPs. In simple terms, the system provides subject matter experts with timely access to data for monitoring and analysis, without incurring the logistical and technical costs associated with manual data collection.

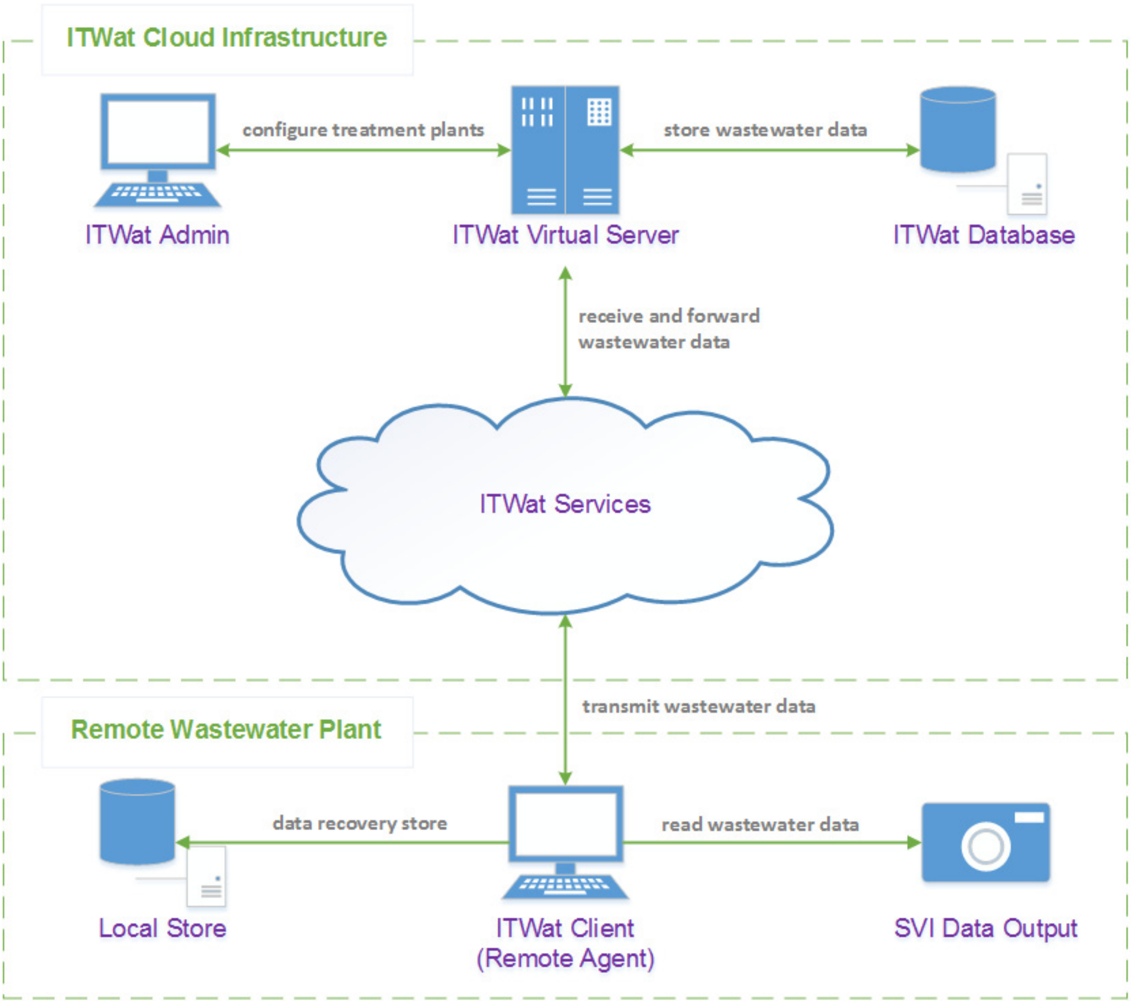

Figure 1 ITWat architecture and components

3.1 ITWat Client

Purpose: The ITWat Client is a background application that runs at set intervals on a local PC in the plant. Its primary responsibility is to collaborate with the cloud-based ITWat Services to determine and execute a strategy for transferring data that resides in the plant, such as the SVI Data Output repository.

Functions: The ITWat Client background application possesses the following functionality;

1. Determine the current available bandwidth for the facility in which it resides.
2. Request a data collection instructions from ITWat Services – the instructions contain a list of data points that need to be read, along with their location, type and resolution/frequency.
3. Parse and store collected data;
    a. When an active connection is available send to ITWat Services.
    b. When an active connection is not available send to the Local Store.

3.2 Local Store

Purpose: The Local Store is a local database that operates as a redundant and alternative storage repository for data. Its primary role in the system is to facilitate the storage of data from ITWat Client when there is no external connection available for transmitting data to the cloud.

Functions: The functionality of the Local Store is basic, and limited to read and write operations;

1. Write data to the repository.
2. Read data within a particular date range from the repository.

3.3 SVI Data Output

Purpose: The SVI Data Output component is a sub-system that provides the SVI values for samples taken in the plant. At present, this is the main type of data being consumed by the data collection system.

Functions: The functionality of the SVI Data Output component is too detailed to address in this paper, but its operation in the context of this paper is summarised below;

1. Calculate SVI values from samples.
2. Encode data using a naming convention to enable integration with ITWat Client.
3. Log SVI values in a local repository for ITWat Client to consume.

3.4 ITWat Services

Purpose: The ITWat Services are a collection of cloud-based services that are responsible for liaising with ITWat Client components that reside in decentralised WWTPs. Through cohesive and synchonrised collaboration, these components work together to realise intelligent, customised and optimised data collection across these plants.

Functions: ITWat Services is the central component in the architecture, and is responsible for bridging communications between local and cloud infrastructures, its primary set of functions include;

1. Listening for communication requests from ITWat Client components across multiple decentralised WWTPs.
2. Dynamically create data collection assignments (e.g. quantity of data to upload) for each plant based on its bandwidth classification.
3. Accepting data transmissions from ITWat Client applications and persisting this data in the cloud for analysis.

3.5 ITWat Virtual Server

Purpose: The ITWat Virtual Server provides the hardware platform on which ITWat Admin and ITWat Services are hosted. Given its passive nature in the system architecture, there are no significant functions that contribute to the data collection cycle.

3.6 ITWat Database

Purpose: The ITWat Database is a cloud-based relational database that is used by ITWat Services to store meta-data about all decentralised WWTPs, as well as persisting data that is transmitted from ITWat Client applications deployed on these sites. The ITWat Database is a multi-tenancy database, which means that a single database can be used to store data across multiple sites.

Functions: The functionality of the ITWat Database is limited to read and write operations that relate to site metadata and collected measurements;

1. Read/write WWTP information that is needed to inform data collection undertaken by ITWat Client, such as the location of data points that need to be interrogated.
2. Read/write data sent to ITWat Services by the local ITWat Client.
3. Return data for a particular WWTP for analysis using a date range.

3.7 ITWat Admin

Purpose: ITWat Admin is a web-based user interface that allows administrators to administrate the ITWat system using a central and accessible tool. By enabling system configuration at the cloud-level, changes in operation can be disseminated dynamically to other components in the system in a timely manner, without the need to recompile or redeploy individual components.

Functions: The functionality of ITWat Admin centres on the configuration of system and site settings. These include;

1. Managing the list of decentralised WWTPs being serviced by the system.
2. Setting default bandwidth profiles to each WWTP to match the connectivity observed on each site – however it should be noted that ITWat Services dynamically changes this based on a sites data transmission history.
3. Managing the list of data points for each WWTP, including the location, type and resolution/frequency of each data source.
4. Visualise the data uploaded from each WWTP, with graphical trends for numerical data, and time-stamped repositories for images.

4 Conclusions

There are several well-known benefits that can be attributed to decentralised WWTPs, such as reduced costs relating to infrastructure, improved water quality and availability, and reduced conventional pollutants and emerging contaminants. However, given the geographically distributed and isolated nature of many of these plants, they can prove difficult to monitor in near-time. Many of the solutions that are currently used for remote monitoring of decentralised WWTPs suffer from being too expensive, delivering highly granular data that is not suitable for timely decision-making, and most importantly, are not resilient to operating in low-bandwidth environments, with intermittent and highly volatile connectivity.

Our distributed system architecture for collecting data from decentralised WWTPs, which was presented in this paper, provides a resilient and reliable solution for transmitting data from these facilities, thereby enabling the timely analysis and monitoring of multiple plants through a single browser-based application. Furthermore, while the intelligent and optimised transmission of data is arguably the most important aspect of our solution, the application of open standards, modular and extensible design, and the ability to disseminate data collection instructions from the cloud without onsite redeployment, all contribute to ITWat being a high-performance and low-maintenance solution for simultaneously monitoring multiple decentralised WWTPs.

References

[1] EU, "Water Framework Directive," 2015. .

[2] Environmental Protection Agency, "STRIVE Innovation for a green economy," 2009.

[3] M. a. Massoud, A. Tarhini, and J. a. Nasr, "Decentralized approaches to wastewater treatment and management: Applicability in developing countries," *J. Environ. Manage.*, vol. 90, no. 1, pp. 652–659, 2009.

[4] US EPA Decentralized Wastewater Memorandum of Understanding Partnership, "Decentralized wastewater treatment: A Sensible Solution." 2005.

[5] N. Brunner and M. Starkl, "Financial and economic determinants of collective action: The case of wastewater management," *Environ. Impact Assess. Rev.*, vol. 32, no. 1, pp. 140–150, 2012.

[6] W.-Y. Chung and J.-H. Yoo, "Remote water quality monitoring in wide area," *Sensors Actuators B Chem.*, vol. 217, pp. 51–57, 2015.

[7] E. Curry, V. Degeler, E. Clifford, D. Coakley, A. Costa, S. J. van Andel, N. van de Giesen, C. Kouroupetroglou, T. Messervey, J. Mink, and S. Smit, "Linked Water Data for Water Information Management," in *11th International Conference on Hydroinformatics (HIC)*, 2014.